Available online at www.sciencedirect.com

ScienceDirect

Theoretical Computer Science 385 (2007) 1–10

Theoretical Computer Science

www.elsevier.com/locate/tcs

On the minimum monochromatic or multicolored subgraph partition problems☆

Xueliang Li[a], Xiaoyan Zhang[b,*]

[a] *Center for Combinatorics and LPMC, Nankai University, Tianjin 300071, PR China*

[b] *School of Mathematics and Computer Science and Institute of Mathematics, Nanjing Normal University, Nanjing 210097, PR China*

Received 1 December 2006; received in revised form 23 April 2007; accepted 29 April 2007

Communicated by D.-Z. Du

Abstract

Let $G = (V, E)$ be an edge-colored graph. A subgraph H is said to be *monochromatic* if all the edges of H have the same color, and *multicolored* if no two edges of H have the same color. We investigate the complexity of the problems for finding the minimum number of monochromatic or multicolored subgraphs, such as cliques, cycles, trees and paths, partitioning $V(G)$, depending on the number of colors used and the maximal number of times a color appears in a coloring. We also present a greedy scheme that yields a $(\ln m + 1)$-approximation for the problem of finding the minimum number of monochromatic cliques partitioning $V(G)$ for a K_4^--free graph G, where m is the size of the largest monochromatic clique in G. By a slightly modification of the approximation algorithm, it can be used for the multicolored case. We show that unless $NP \subseteq DTIME(\mathrm{N}^{\mathrm{O}(\log\log \mathrm{N})})$, for any $\epsilon \geq 0$ there is no approximation algorithm for finding the minimum number of multicolored trees partitioning $V(G)$ with performance $50/521(1-\epsilon)\ln |V|$.

Keywords: Monochromatic; Multicolored; Partition; Complexity; Inapproximability

1. Introduction

For graphs, people often assigned certain weights, labels or colors representing certain types or costs of the relations between the vertices. In many of these problems, the goal is to find some subgraphs in which a function of weights, labels or colors of the edges attains some optimum value. Well-known examples are the Minimum Spanning Tree problem and the Traveling Salesman problem. Other examples are problems in which one is interested in monochromatic cliques (trees, cycles or paths), i.e., cliques (trees, cycles or paths) in which all edges have the same color, or multicolored cliques (trees, cycles or paths), i.e., cliques (trees cycles or paths) in which all edges have different colors.

☆ Supported by NSFC, PCSIRT, the “973” program, and Jiangsu Planned Projects for Postdoctoral Research Funds (0602023C).

* Corresponding author. Tel.: +86 138 2098 5339.
E-mail addresses: lxl@nankai.edu.cn (X. Li), xiaoyanice@yahoo.com.cn (X. Zhang).

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2007.04.033

Manoussakis et al. [23] presented conditions on the minimum number k of colors sufficient for the existence of properly edge-colored subgraphs in a k-edge colored complete graphs K_n. The types of subgraphs they studied include families of internally pairwise vertex-disjoint paths with common endpoints, hamiltonian paths and hamiltonian cycles, cycles with a given lower bound of their length, spanning trees, stars, and cliques. Erdős et al. [8] showed that if the edges of a finite complete graph are colored with r colors, then the vertex set of the graph can be covered by at most $cr^2 \log r$ vertex-disjoint monochromatic cycles, where c is a constant. The analogue problems have been extensively studied recently, see Haxell and Kohayakawa [16], Haxell [17], Kaneko et al. [20] and Jin et al. [18]. The multicolored partition problems and the related problems were studied by Alon [2], Erdős and Tuza [9], Albert et al. [1], Brualdi and Hollingsworth [6], Kaneko et al. [20]. Broersma et al. [5] gave some basic results on paths and cycles in general colored graphs.

Most of these colored problems can be viewed as special cases of graph partitions, and one can observe several variations and interesting aspects as far as the problem of coloring and partition is concerned. Feder et al. [11] introduced a parameterized family of graph problems which contain some well-known graph partition problems as special cases. Alon et al. [3] presented several results on edge partitions and vertex partitions of graphs into graphs with components of bounded size. MacGillivray and Yu [22] studied the (H, C)-partition problem which is a general graph partition problem containing some well-known graph partition problems as special cases. Yegnanarayanan [24] considered three coloring parameters of a graph G in connection with the computational complexity, partitions, algebra, projective plane geometry and analysis. Many graph partition problems with their corresponding computational complexity [4,7,11,13,15,19] have been well studied.

By a coloring of G we mean a surjective function $l : E \to \{1, 2, \ldots, r\}$. If G is assigned such a coloring, we say that G is an *edge-colored graph*. We call $l(e)$ the color of the edge $e \in E$, and use $l(H)$ to denote the number of different colors in the set $\{l(e) \mid e \in E(H)\}$ for a subgraph H of G. Let $E_i = \{e \in E(G) \mid l(e) = i\}$ for $i = 1, 2, \ldots, r$. We use $s(G)$ to denote the number $max\{|E_i| \mid i = 1, 2, \ldots, r\}$. A clique C of G is called a *monochromatic clique* if and only if all edges in C have the same color. A clique C of G is called a *multicolored-clique* if and only if no two edges in C have the same color. *Monochromatic cycles, trees or paths*, and *multicolored cycles, trees or paths* can be defined similarly. The *Minimum Monochromatic Clique Partition* problem or the *Minimum Multicolored Clique Partition* problem is to find a partition of $V(G)$ into minimum number of monochromatic cliques or multicolored cliques, respectively. The *Minimum Monochromatic Cycle, Tree or Path Partition* problem and the *Minimum Multicolored Cycle, Tree or Path Partition* problem can be defined similarly. Note that a single vertex is also regarded as a monochromatic or multicolored clique, cycle, tree or path, which is simply called a vertex-monochromatic or vertex-multicolored clique, cycle, tree or path, respectively. In this paper, we investigate the complexity of partitioning the vertex set of an edge-colored graph into the monochromatic or multicolored subgraphs, such as cliques, cycles, trees or paths, depending on the parameters $s(G)$ and $l(G)$. Let us consider the following questions first.

1. The Minimum Monochromatic Clique Partition problem or the Minimum Multicolored Clique Partition problem is equivalent to the edge cover problem which can be solved in polynomial time by graph matching algorithm [13, 21] when $s(G) \leq 2$ or $l(G) \leq 2$, respectively. Then, we ask what is the complexity of the Minimum Monochromatic Clique Partition problem or the Minimum Multicolored Clique Partition problem when $s(G) = 3$ or $l(G) = 3$, respectively. Can the problem be solved in polynomial time? Or, is $s(G) = 3$ or $l(G) = 3$ the threshold when the problem becomes NP-complete?

2. The Minimum Monochromatic Cycle Partition or the Minimum Multicolored Cycle Partition of G is the n number of vertex-monochromatic cycles or vertex-multicolored cycles which are the n vertices of G when $s(G) \leq 2$ or $l(G) \leq 2$, respectively. Whereas, when G is monochromatic or multicolored, finding Minimum Monochromatic Cycle Partition or Minimum Multicolored Cycle Partition is somehow equivalent to finding a hamiltonian cycle of G. It is well known that the Hamiltonian Cycle problem is *NP*-complete. Then, we ask what is the complexity of Minimum Monochromatic Cycle Partition problem or the Minimum Multicolored Cycle Partition problem when $s(G) = 3$ or $l(G) = 3$, respectively. Can the problem be solved in polynomial time? Or, is $s(G) = 3$ or $l(G) = 3$ the threshold when the problem becomes NP-complete?

3. At first, with simple observation the situation for the Minimum Monochromatic Path Partition problem or the Minimum Multicolored Path Partition problem is the same as that for the Minimum Monochromatic Tree Partition problem or the Minimum Multicolored Tree Partition problem when $s(G) = 1$ or $l(G) = 1$, respectively. All of the problems are equivalent to the edge cover problem and can be solved in polynomial time by graph matching algorithm

[13,21]. It is easy to see that the situations are also the same for the Minimum Monochromatic Path Partition problem and the Minimum Monochromatic Tree Partition problem or the Minimum Multicolored Path Partition problem and the Minimum Multicolored Tree Partition problem when $s(G) = 2$ or $l(G) = 2$. However, the situations are different when $s(G) = |E(G)|$ or $l(G) = |E(G)|$. In this case, finding the Minimum Monochromatic Path Partition or the Minimum Multicolored Path Partition of G is somehow equivalent to finding a hamiltonian path of G, and therefore is NP-complete, whereas, finding the Minimum Monochromatic Tree Partition or the Minimum Multicolored Tree Partition of G is equivalent to finding a spanning tree of G which can also be solved in polynomial time. Then, it is interesting for us to consider the complexity of the Minimum Monochromatic Path (Tree) Partition problem or the Minimum Multicolored Path (Tree) Partition problem when $s(G) = 2$ or $l(G) = 2$, respectively.

This paper is organized as follows. In Section 2, we show that the Minimum Monochromatic Clique Partition problem is NP-complete even if the input graph G is K_4^--free and $s(G) = 3$. We also present a greedy scheme that yields a $\ln m + 1$-approximation for the problem by a general color assignment, where m is the size of the largest monochromatic cliques. In Section 3, we utilize the simple technique introduced in Section 2 to investigate all the other problems and answer the questions proposed in Section 1 completely. In Section 4, we derive the inapproximability result for the Minimum Multicolored Tree Partition problem for a general color assignment. In Section 5, we propose some problems for further study and give concluding remarks on some simple observations of the approximation factor for the Minimum Multicolored Tree and Path Partition problems.

2. Minimum monochromatic clique partition

In this section, we will use the Exact Cover By 3-Sets(X3C) problem to show that the Minimum Monochromatic Clique Partition problem is *NP*-complete even if the graph is K_4^--free and $s(G) = 3$. This technique can be applied easily to obtain the complexity of the other problems in the following section. We present a $\ln m + 1$-approximation algorithm for the problem by a general color assignment, where m is the size of the largest monochromatic clique in G. First, some additional terminology is needed.

A clique in a graph is a complete subgraph, and a clique of size i is denoted as K_i. Note that if two cliques share an edge then both cliques are of size at least 3. A graph G is called K_4^--free if it does not contain an induced K_4^- as a subgraph, where K_4^- is obtained by deleting one edge from K_4. A vertex x is *color-adjacent* to a vertex y of a monochromatic clique C in G if xy is an edge in G with the same color as that of C. Given an edge-colored graph $G = (V, E)$, the number of monochromatic cliques in a monochromatic clique partition P^C is denoted by $|P^C(G)|$. Next, we give a formal description for the Minimum Monochromatic Clique Partition problem as follows.

Minimum Monochromatic Clique Partition problem

INSTANCE: A graph $G = (V, E)$, a coloring $l : E \to \mathbb{N}$, and a positive integer $k \le n$.

QUESTION: Is there a monochromatic clique partition P^C of G with $|P^C(G)| \le k$?

According to [13], we know that the following problem is *NP*-complete.

Exact Cover By 3-Sets(X3C) problem

INSTANCE: A set X with $|X| = 3q$, and a collection S of 3-element subsets of X.

QUESTION: Does S contain an exact cover for X, i.e., a subcollection $S^* \subseteq S$ such that every element of X occurs in exactly one member of S^* ?

Theorem 2.1. *The Minimum Monochromatic Clique Partition problem is NP-complete even if the input graph G is K_4^--free and $s(G) = 3$.*

Proof. Clearly, the problem is in *NP*. Guess a set of k cliques of size at most 3 and check in polynomial time that whether all the cliques in the set are vertex-disjoint monochromatic ones which cover all the vertices of the given edge-colored graph.

To complete the proof, we shall reduce the Exact Cover By 3-Sets problem to the Minimum Monochromatic Clique Partition problem. Let an arbitrary instance of the Exact Cover By 3-Sets problem be given by the set $X = \{x_1, \dots, x_{3q}\}$ and a collection $S = \{s_1, \dots, s_m\}$ of 3-element subsets of X. We construct a K_4^--free graph G with every color appearing at most 3 times. Start with the collection $S = \{s_1, \dots, s_m\}$ and for every 3 vertices x_{i1}, x_{i2}, x_{i3} such that $s_i = \{x_{i1}, x_{i2}, x_{i3}\}$, $i = 1, \dots, m$, we add a gadget H_i which consists of vertices x_{i1}, x_{i2}, x_{i3} and 6 new vertices $y_{i1}, y_{i2}, y_{i3}, z_{i1}, z_{i2}, z_{i3}$ with the following color edges: edge $x_{i1}y_{i1}$, edge $y_{i1}z_{i1}$ and edge $x_{i1}z_{i1}$ labeled with color l_{i1}; edge $x_{i2}y_{i2}$, edge $y_{i2}z_{i2}$ and edge $x_{i2}z_{i2}$ labeled with color l_{i2}; edge $x_{i3}y_{i3}$, edge $y_{i3}z_{i3}$, and edge $x_{i3}z_{i3}$ labeled with color l_{i3}; edge $z_{i1}z_{i2}$, edge $z_{i2}z_{i3}$ and edge $z_{i1}z_{i3}$ labeled with color l_{i4}; edge $y_{i1}y_{i2}$, edge $y_{i2}y_{i3}$ and edge $y_{i1}y_{i3}$ labeled with color l_{i5}. It is easy to check that G is K_4^--free and $s(G) = 3$, which contains $3q + 6m$ vertices and the least possible number of monochromatic cliques partitioning $V(G)$ should be $(3q + 6m)/3 = q + 2m$ since every monochromatic clique has at most 3 vertices. Clearly the construction can be accomplished in polynomial time. We will show that there exists a subcollection $S^* \subseteq S$ with q subsets covering all the elements of X such that every elements of X occurs in exactly one member of S^* if and only if there are the least possible $q + 2m$ monochromatic vertex-disjoint cliques which cover all the vertices of G, i.e., we define the positive integer k to be $q + 2m$.

(The "if" part) suppose there exists a subcollection $S^* = \{s_1^*, \dots, s_q^*\}$ which exactly covers all the $3q$ elements of X, i.e., $s_i^* \cap s_j^* = \emptyset$, $i \neq j$, $i, j \in [1, q]$.

For each $s_i^* = \{x_{i1}^*, x_{i2}^*, x_{i3}^*\}$, $i \in [1, q]$. We choose 3 vertex-disjoint monochromatic cliques: clique $z_{i1}^* y_{i1}^* x_{i1}^*$, clique $z_{i2}^* y_{i2}^* x_{i2}^*$ and clique $z_{i3}^* y_{i3}^* x_{i3}^*$ to cover the set of 9 vertices of H_i. Then all the vertices of $\{x_i | i = 1, \dots, 3q\}$ have been covered by these monochromatic cliques. Denote $S \setminus S^*$ by S'. Clearly $|S'| = m - q$. For each s_i' in S', we choose 2 vertex-disjoint monochromatic cliques: clique $z_{i1}' z_{i2}' z_{i3}'$ and clique $y_{i1}' y_{i2}' y_{i3}'$ to cover the left uncovered vertices of G. Thus we totally obtain $3q + 2(m - q) = 2m + q$ vertex-disjoint monochromatic cliques which cover all the $3q + 6m$ vertices of G.

(The "only if" part) Suppose there are $q + 2m$ vertex-disjoint monochromatic cliques which cover all the vertices of G. Let C denote the set of all the $q + 2m$ cliques. Then every monochromatic clique in C must exactly consist of 3 vertices since $q + 2m$ is the least possible number of vertex-disjoint monochromatic cliques which can cover all the $3q + 6m$ vertices of G. Let us make an observation: there are only two kinds of ways for the monochromatic cliques to exist in each gadget H_i. The first kind of way is to choose clique $z_{i1}y_{i1}x_{i1}$, clique $z_{i2}y_{i2}x_{i2}$ and clique $z_{i3}y_{i3}x_{i3}$; The second kind of way is to choose clique $z_{i1}z_{i2}z_{i3}$ and clique $y_{i1}y_{i2}y_{i3}$. Otherwise, if the monochromatic cliques are not found in the above two kinds of ways for each H_i, it is easy to verify that there will exist some vertices in H_i which can not be covered by any 3-vertex monochromatic clique. Further, if the monochromatic cliques are found in the first kind of way for H_i, $i = 1, \dots, m$, the corresponding s_i will be added into S^*. Since all vertex-disjoint monochromatic cliques which are found in the first kind of way can cover all the vertices of X, all the elements of X can be covered by S^* such that every element of X occurs in exactly one member of S^*. Since the total number of x_i is $3q$, the number of s_i in S^* must be q. This completes the proof. □

The *Maximum Clique* problem is to find a clique of the largest size in a graph G. It is a well-known problem that the maximum clique problem is not only NP-complete but also very difficult to approximate. Hastad [14] proved that the Maximum Clique problem is hard to approximate within $n^{1-\epsilon}$ for any $\epsilon > 0$. However, we can find a largest monochromatic clique in polynomial time when the input graph G is K_4^--free. We give the algorithm as follows:
The algorithm of finding a largest monochromatic clique:
1. Input G and let $C := \phi$;
2. Repeat: start from any edge $v_i v_j \in E(G)$,
   Let $S := \phi$, $S := S \cup \{v_i\} \cup \{v_j\}$;
   While there is a vertex v_k which is color-adjacent to each vertex of S
      Do $S := S \cup \{v_k\}$;
   End While
   Let $C := C \cup S$, $E(G) := E(G) - E(S)$;
 Until no edge in $E(G)$.
3. Return the largest set in C.

Lemma 2.2. *The above algorithm can output a largest monochromatic clique for any K_4^--free graph G in polynomial time.*

Proof. We claim that any two cliques do not share an edge in a K_4^--free graph. Otherwise suppose that clique C_1 and clique C_2 share an edge e which is incident with vertices i and j. Then, there must exist two vertices $a \in V(C_1)$ and $b \in V(C_2)$ such that the subgraph $aibj$ forms a K_4^--free subgraph, a contradiction and the claim holds. This implies that every edge just belongs to one clique. The above algorithm starts from any edge e and find all the same color edges in the monochromatic clique which contains edge e and then delete all the edges in the clique. So the algorithm will find all the monochromatic cliques after the edge set of G becomes empty. The process takes at most $O(|E||V|^2) = O(|V|^4)$ time. We can easily choose a largest monochromatic clique from the set of monochromatic cliques found by the algorithm. □

By using the algorithm of finding a largest monochromatic clique, we can present the greedy algorithm to solve the Minimum Monochromatic Clique Partition problem for a K_4^--free graph G by a general color assignment.
Greedy algorithm for the Minimum Monochromatic Clique Partition problem:
1. Input G and let $CP := \phi$;
2. While the vertex set of G is not empty
    $C :=$ a largest monochromatic clique found in G;
    Delete G by C, i.e., $G := G \setminus C$;
    $CP := CP \cup C$;
  End While
3. Return the set of monochromatic cliques in CP.

Theorem 2.3. *The above algorithm can achieve the performance ratio at most* $\ln m + 1$, *where m is the size of the largest monochromatic clique.*

Proof. Let $G = (V, E)$ be a K_4^--free graph and m is the size of the largest monochromatic clique in G. Denote by C_i the set of monochromatic cliques of size i, $1 \le i \le m$, obtained by executing the greedy algorithm. Let n_i be the number of monochromatic cliques in C_i. Let $S = \cup_{i=1}^{m} C_i$ and then $|S| = \sum_{i=1}^{m} n_i$, i.e., $|S|$ is the partition number obtained by the greedy algorithm. Since the number of vertices of all the monochromatic cliques in S is equal to $|V|$, we have

$$mn_m + (m-1)n_{m-1} + \cdots + 2n_2 + n_1 = |V|. \tag{1}$$

Denote S_{opt} as the set of monochromatic cliques which partitions V into minimum number of vertex-disjoint monochromatic cliques. Let $|S_{opt}|$ be the partition number of S_{opt}.

First, we can easily observe that

$$n_1 \le |S_{opt}|. \tag{2}$$

We have denoted C_1 as the set of monochromatic cliques of size 1 in S which are the vertex-monochromatic cliques. And we have $|C_1| = n_1$. From the greedy rule of the algorithm, we know that there does not exist an edge between any two vertex-monochromatic cliques in G. Hence, any two vertex-monochromatic cliques in C can not lie in a same monochromatic clique in S_{opt} which implies $n_1 \le |S_{opt}|$.

Let $\Delta_1 = (|S_{opt}| - n_1)/2$, $\overline{n}_2 = n_2 - \Delta_1$ and $n_1' = n_1 + 2\Delta_1 = |S_{opt}|$. Then, we can get from (1) that

$$mn_m + (m-1)n_{m-1} + \cdots + 2\overline{n}_2 + n_1' = |V|. \tag{3}$$

Let $|S_1| = \sum_{i=3}^{m} n_i + \overline{n}_2 + n_1'$. Since $\Delta_1 \ge 0$ from (2), we have

$$|S| = \sum_{i=1}^{m} n_i \le \sum_{i=1}^{m} n_i + \Delta_1 = \sum_{i=3}^{m} n_i + \overline{n}_2 + n_1' = |S_1|.$$

Next, we claim that

$$2|S_{opt}| \ge 2n_2 + n_1.$$

We delete S_{opt} by $C_m \cup \cdots \cup C_3$ to obtain a new set of monochromatic cliques S'_{opt}. Then the number of monochromatic cliques in S'_{opt} is at most $|S_{opt}|$. It is easily seen that according to the greedy rule of the algorithm, every monochromatic clique in S'_{opt} has size at most 2. Since $|V(S_{opt})| = |V|$, the number of vertices of all the monochromatic cliques in S'_{opt} is $|V(S_{opt})| - \sum_{i=3}^{m} |V(C_i)| = |V| - \sum_{i=3}^{m} |V(C_i)| = 2n_2 + n_1$. This implies that $2|S_{opt}| \geq 2n_2 + n_1$. So the claim holds.

From Eq. (1), we have

$$mn_m + (m-1)n_{m-1} + \cdots + 3n_3 \geq |V| - 2|S_{opt}|.$$

We, therefore, have $2\overline{n}_2 + n'_1 = 2\overline{n}_2 + |S_{opt}| \leq 2|S_{opt}|$ from (3). This implies

$$2\overline{n}_2 \leq |S_{opt}|. \tag{4}$$

Let $\Delta_2 = (|S_{opt}| - 2\overline{n}_2)/3$, $\overline{n}_3 = n_3 - \Delta_2$ and $n'_2 = \overline{n}_2 + 3\Delta_2/2 = |S_{opt}|/2$. We get from (3) that

$$mn_m + (m-1)n_{m-1} + \cdots + 3\overline{n}_3 + 2n'_2 + n'_1 = |V|.$$

Let $S_2 = \sum_{i=4}^{m} n_i + \overline{n}_3 + \sum_{i=1}^{2} n'_i$. Then since $\Delta_2 \geq 0$ from (4), we get

$$|S| \leq |S_1| \leq \sum_{i=3}^{m} n_i + \overline{n}_2 + n'_1 + 1/2\Delta_2 = \sum_{i=4}^{m} n_i + \overline{n}_3 + \sum_{i=1}^{2} n'_i = |S_2|.$$

Similarly, we can continue the strategy to prove that $i|S_{opt}| \geq in_i + (i-1)n_{i-1} + \cdots + n_1$, $i = 3, \ldots, m-1$, and keep the equality

$$mn_m + (m-1)n_{m-1} + \cdots + i\overline{n}_i + (i-1)n'_{i-1} + \cdots + n'_1 = |V| \tag{5}$$

where $n'_j = |S_{opt}|/j$, $1 \leq j \leq i-1$.

From (1), we have

$$mn_m + (m-1)n_{m-1} + \cdots + (i+1)n_{i+1} \geq |V| - i|S_{opt}|$$

which yields from (5)

$$i\overline{n}_i \leq |S_{opt}|. \tag{6}$$

Let $\Delta_i = (|S_{opt}| - i\overline{n}_i)/(i+1)$, $\overline{n}_{i+1} = n_{i+1} - \Delta_i$ and $n'_i = \overline{n}_i + (i+1)\Delta_i/i = |S_{opt}|/i$. Then equality

$$mn_m + (m-1)n_{m-1} + \cdots + (i+1)\overline{n}_{i+1} + in'_i + \cdots + n'_1 = |V| \tag{7}$$

holds where $i = 3, \ldots, m-1$.

Let $|S_i| = \sum_{j=i+2}^{m} n_j + \overline{n}_{i+1} + \sum_{j=1}^{i} n'_j$. Then

$$|S| \leq |S_1| \leq \cdots \leq |S_{i-1}| \leq |S_{i-1}| + \Delta_i/i = |S_i|$$

since $\Delta_i \geq 0$ from (6) where $i = 3, \ldots, m-1$.

Finally, when $i = m$, we have got the equality from (7)

$$m\overline{n}_m + (m-1)n'_{m-1} + \cdots + n'_1 = |V| \tag{8}$$

where $n'_j = |S_{opt}|/j$, $1 \leq j \leq m-1$.

With the obvious fact that $m|S_{opt}| \geq |V|$ since every monochromatic clique in S_{opt} has size at most m, we have from (8)

$$m\overline{n}_m = |V| - (m-1)|S_{opt}| \leq m|S_{opt}| - (m-1)|S_{opt}| = |S_{opt}|.$$

Then $\overline{n}_m \leq |S_{opt}|/m$. Now let $n'_m = |S_{opt}|/m$. Since

$$|S| \leq |S_1| \leq \cdots \leq |S_{m-1}| = \overline{n}_m + \sum_{j=1}^{m-1} n'_j \leq n'_m + \sum_{j=1}^{m-1} n'_j = \sum_{j=1}^{m} n'_j$$

where $n'_j = |S_{opt}|/j$, $1 \leq j \leq m$, we have $|S| \leq (\sum_{j=1}^{m} 1/j)|S_{opt}| = \mathcal{H}_m|S_{opt}| \leq (1 + \ln m)|S_{opt}|$ where $\mathcal{H}_m$ is the m-th harmonic number. This indicates that the partition number of vertex-disjoint monochromatic cliques obtained from the greedy algorithm is at most $1 + \ln m$ times that of the optimal solution. The theorem holds. □

By a slight modification of the approximation algorithm, we can solve the Minimum Multicolored Clique Partition problem for a K_4^--free graph.

3. The complexity of the other problems

Actually, the technique used to prove that Minimum Monochromatic Clique Partition problem is NP-complete can be easily applied to show the complexity of the Minimum Multicolored Clique, Minimum Monochromatic or Multicolored Cycle, Tree and Path Partition Problems.

For the Minimum Multicolored Clique Partition problem for the graph G with $l(G) = 3$, we can set the gadget H_i to be the same as that in Theorem 2.1 only with different color assignment. We can set the gadget H_i which consists of vertices x_{i1}, x_{i2}, x_{i3} and 6 new vertices $y_{i1}, y_{i2}, y_{i3}, z_{i1}, z_{i2}, z_{i3}$ with the following colored edges: edge $x_{i1}y_{i1}$ labelled with color l_1, edge $y_{i1}z_{i1}$ labelled with color l_2 and edge $x_{i1}z_{i1}$ labelled with color l_3; edge $x_{i2}y_{i2}$ labelled with color l_1, edge $y_{i2}z_{i2}$ labelled with color l_2 and edge $x_{i2}z_{i2}$ labelled with color l_3; edge $x_{i3}y_{i3}$ labelled with color l_1, edge $y_{i3}z_{i3}$ labelled with color l_2, $x_{i3}z_{i3}$ labelled with color l_3; edge $z_{i1}z_{i2}$ labelled with color l_1, edge $z_{i2}z_{i3}$ labelled with color l_2 and edge $z_{i1}z_{i3}$ labelled with color l_3; edge $y_{i1}y_{i2}$ labelled with color l_1, edge $y_{i2}y_{i3}$ labelled with color l_2 and edge $y_{i1}y_{i3}$ labelled with color l_3. Then the other process of the proof is similar to that in the proof of Theorem 2.1. So the Minimum Multicolored Clique Partition problem is NP-complete when there are at most 3 colors even if the graph is K_4-free, while for the graphs with edge-colored by less than 3 colors, it can be solved in polynomial time.

Since clique K_3 is just a cycle of 3 vertices, when every color occurs at most 3 times or there are at most 3 colors, we can use the same proof of the Minimum Monochromatic or Multicolored Clique Partition problem to prove that the Minimum Monochromatic or Multicolored Cycle Partition problem is NP-complete, respectively.

For the Minimum Monochromatic or Multicolored Tree Partition problem, when every color occurs at most twice or there are at most 2 colors, we can set the gadget H_i to be similar to that in Theorem 2.1 or the gadget H_i to be similar to that for proving the complexity of the Minimum Multicolored Clique Partition problem only deleting the edges of $x_{i1}z_{i1}$, $x_{i2}z_{i2}$, $x_{i3}z_{i3}$, $z_{i1}z_{i3}$ and $y_{i1}y_{i3}$. The other process of the proof is similar to that in the proof of Theorem 2.1. Then the Minimum Monochromatic or Multicolored Tree Partition problem is NP-complete when $s(G) = 2$ or $l(G) = 2$, while in the cases when $s(G) = 1$ and $s(G) = n$ or $l(G) = 1$ and $l(G) = n$, it can be solved in polynomial time, respectively.

When every color occurs at most twice or there are at most 2 colors, it is easy to see that the situation for the Minimum Monochromatic or Multicolored Path Partition Problem is identical to that for the Minimum Monochromatic or Multicolored Tree Partition problem. So both of the problems are also NP-complete while the Minimum Monochromatic or Multicolored Path Partition Problem can be solved in polynomial time when $s(G) = 1$ or $l(G) = 1$, respectively.

To summarize, we have answered the questions proposed in Section 1 completely.

4. Minimum multicolored tree partition

We have shown that the Minimum Multicolored Tree Partition problem is NP-complete even for the graphs with edge-colored by 2 colors. Next, for a general color assignment on the edges of G, we achieve an inapproximabily result of this problem. Based on the reduction proof, we can also draw a conclusion that this problem remains to be *NP*-complete for bipartite graphs. The results will be shown in the sequel. First we need some additional terminology.

Given a graph $G = (V, E)$, the number of multicolored trees in a multicolored tree partition P^T is denoted by $|P^T(G)|$.

Minimum Multicolored Tree Partition problem

INSTANCE: A graph $G = (V, E)$, a coloring $l : E \to \mathbb{N}$, and a positive integer $k \leq n$.

QUESTION: Is there a multicolored tree partition P^T of G with $|P^T(G)| \leq k$?

The *Minimum Set Cover* problem can be informally stated as follows:

Minimum Set Cover problem

INSTANCE: A universe U of n elements, a collection of subsets of U, $S = \{s_1, \ldots, s_m\}$, a cost function $c : S \to Q^+$ and a positive integer $k \le \min\{m, n\}$.

QUESTION: Is there a subcollection C of S with $c(C) \le k$ that covers all the elements of U?

Theorem 4.1. *Unless NP $\subseteq$ DTIME*$(\mathrm{N}^{\mathrm{O}(\log\log \mathrm{N})})$*, for any $\epsilon > 0$, there is no approximation algorithm for the Minimum Multicolored Tree Partition problem with performance* $50/521(1-\epsilon)\ln|V|$.

Proof. To present the inapproximability result for the problem, we need to transform any instance of the Minimum Set Cover problem to an instance of the Minimum Multicolored Tree Partition problem first. We construct an edge-colored graph G such that there is a covering $C \subseteq S$ of U with no more than k subsets if and only if G contains $k+1$ or less vertex-disjoint multicolored trees that cover all the vertices of G.

The graph G is constructed as follows: The set of vertices of G is the union of the sets $U = \{u_1, \ldots, u_n\}$, $U' = \{u'_1, \ldots, u'_n\}$ and $S = \{s_1, \ldots, s_n\}$ with one more special vertex v_0.

The set of edges of G with colors is defined as follows:

1. $u_i s_j$ with color l_i if and only if $u_i \in s_j$, $1 \le i \le n$, $1 \le j \le m$;
2. $v_0 s_j$ with color l_{0j}, $1 \le j \le m$;
3. $v_0 u'_i$ with color l_i, $1 \le i \le n$;

Clearly, the construction can be accomplished in polynomial time.

Without loss of generality, suppose the set of $\{s_j | 1 \le j \le k\}$ covers all the elements of U. Then, it is easy to find k vertex-disjoint multicolored trees with roots s_j, $1 \le j \le k$, respectively, which cover all vertices of U. The left vertices can form a multicolored tree with root v_0. So G contains $k+1$ or less multicolored trees which cover all vertices of G.

Conversely, if there is a multicolored tree partition P^T of G with $k+1$ or less multicolored trees, we can find a subcollection C of S with k or less number of subsets which cover all elements of U.

First, let us consider such a case that there exists a multicolored tree T which contains at least the vertex v_0, some vertices of U and some vertices of S. In such a case, let us think about the smallest number of multicolored trees partitioning $V(G)$ which are obtained from P^T. Suppose there are t_u vertex-multicolored trees of U, denoted by $U^- = \{u_{n-t_u+1}, \ldots, u_n\}$. Without loss of generality, suppose T contains the set of vertices $U^+ = \{u_1, \ldots, u_p\}$ of U, where $p \le n - t_u$, and assume that the left vertices u_i, $p+1 \le i \le n-t_u$ are contained in t multicolored trees rooted at t vertices in S, where $t \le m$. Then the corresponding vertices u'_i, $1 \le i \le p$, in U' cannot be contained in T and become p vertex-multicolored trees, otherwise T would have some common colors. Clearly, when G has the smallest number of multicolored trees, the number of vertex-multicolored trees in G is $p + t_u$, i.e., there is no vertex-multicolored tree formed by a vertex in S. Otherwise, suppose there were some vertex-multicolored trees formed by the vertices in S, then we can merge them into T and decrease the number of multicolored trees partitioning $V(G)$. This causes a contradiction to the smallest property. So the smallest number of multicolored trees partitioning $V(G)$ is $1 + t + p + t_u$. Since $1 + t + p + t_u \le 1 + k$, we have $t + p + t_u \le k$. It is easy to find at most $p + t_u$ subsets of U in S which can cover $U^+ \cup U^-$. All elements in $U - (U^+ \cup U^-)$ can be covered by t subsets of U in S which correspond to the t multicolored trees rooted at t vertices in S. Hence, we can find $t + p + t_u$, i.e., k or fewer subsets of U in S which cover all elements of U.

Second, if there does not exist a multicolored tree T which contains the vertex v_0, some vertices of U and some vertices of S together, then in such a case, let us also consider the smallest number of multicolored trees partitioning $V(G)$ which are obtained from P^T. Suppose the number of vertex-multicolored trees of U is t_u, the number of vertex-multicolored trees of S is t_s and the number of multicolored trees composed of both the vertices in S and the vertices in U is t. It is easy to observe that the number should be at most $t_s + t + t_u + 1$, i.e., all vertices of U' belong to a multicolored tree T' rooted at v_0. While the t_s vertex-multicolored trees of S can be merged into T', then the smallest number is $t + t_u + 1 \le 1 + k$, and hence $t + t_u \le k$. It is easy to find at most t_u subsets of U in S which can cover the t_u elements in U corresponding to the t_u vertex-multicolored trees of U. The other elements of U can be covered by t subsets of U in S corresponding to the t vertices in S which are the t roots of the multicolored trees composed of both the vertices in S and the vertices in U. Thus, we can find a subcollection of S with $t + t_u$, i.e., k or less subsets which can cover all elements of U.

Further, for proving the inapproximability of the Minimum Multicolored Tree problem, we need use a theorem from [12] as follows: Unless $NP \subseteq DTIME$ ($\mathrm{N}^{\mathrm{O}(\log\log \mathrm{N})}$), for any $\epsilon > 0$, there is no approximation algorithm for the Minimum Set Cover problem with performance $(1-\epsilon)\ln|U|$.

The instances in the proof of the above theorem have a special property which has been indicated in [10]. Let (U, S) be any instance produced by the reduction in [12] used for the proof of the above theorem. Then, the special property is $|S| \leq |U|^5$, which means that the number of subsets is polynomially bounded by the number of universal elements. Given an instance $I = (U, S)$ of Minimum Set Cover problem as constructed in [12].

From the construction of G in the proof, it is easy to see that the number of vertices $|V|$ satisfies $|V| = 2|U|+|S|+1$. Since $|S| \leq |U|^5$, we have $|V| \leq 2|U| + |U|^5 + 1$. Let $L(|U|) = 1/(2/|U|^4 + 1 + 1/|U|^5)$, $R(|U|) = |U|^{21/100}$ and $Y(|U|) = (L(|U|))(R(|U|)) = (1/(2/|U|^4 + 1 + 1/|U|^5))(|U|^{21/100})$. Then, since both $L(|U|) = 1/(2/|U|^4 + 1 + 1/|U|^5)$ and $R(|U|) = |U|^{21/100}$ monotonically increase in the range of $|U| \in [2, +\propto)$, and $Y(2) = L(2)R(2) \geq 1$, we can conclude that $Y(|U|) \geq 1$ for $|U| \geq 2$. By adjustment, $Y(|U|) = (|U|^{521/100})/(2|U| + |U|^5 + 1)$, and thus $2|U| + |U|^5 + 1 \leq |U|^{521/100}$ for $|U| \geq 2$, i.e., $|V| \leq |U|^{521/100}$, for $|U| \geq 2$.

Assume that for some constant $\alpha \leq 50/521$, there was an approximation algorithm g for the Minimum Multicolored Tree Partition problem on an edge-colored graph of order $|V|$ with the performance $\alpha(1-\epsilon)\ln|V|$. Let OPT_{tree} denote the optimum value of an instance of the Minimum Multicolored Tree Partition problem, and let OPT_{set} denote the optimum value of an instance of the Minimum Set Cover problem. We have got that there is a covering $C \subseteq S$ of U with no more than k subsets if and only if G contains $k+1$ or less vertex-disjoint multicolored trees that cover all the vertices of G, then $OPT_{tree} = OPT_{set} + 1$.

Since from the set of multicolored trees of any partition P^T, if we remove the multicolored tree which contains the vertex v_0, we can obtain a set cover with size no more than $|P^T| - 1$ according to the left $|P^T| - 1$ multicolored trees from the above argument. Then, when we run g on the graph G constructed in the proof, a solution can be output with at most $\alpha(1-\epsilon)\ln|V| * OPT_{tree}$ multicolored trees, which can be transformed into a solution T_{set} of set cover of at most $\alpha(1-\epsilon)\ln|V| * OPT_{tree} - 1$. And since $|V| \leq |U|^{521/100}$, we have that $\alpha(1-\epsilon)\ln|V| * OPT_{tree} \leq 521/100\alpha(1-\epsilon)\ln|U| * OPT_{tree}$, for $|U| \geq 2$. Hence, $T_{set} \leq 521/100\alpha(1-\epsilon)\ln|U| * OPT_{tree} - 1 = 521/100\alpha(1-\epsilon)\ln|U| * (OPT_{set} + 1) - 1 \leq 521/100\alpha(1-\epsilon)\ln|U| * (OPT_{set} + OPT_{set})$. Moreover, since we have assumed that $\alpha \leq 50/521$, then $T_{set} \leq (1-\epsilon)\ln|U| * OPT_{set}$. This is a contradiction to the inapproximability of the Minimum Set Cover problem. Hence, the assumption does not hold. Then, unless $NP \subseteq DTIME(\mathrm{N}^{\mathrm{O}(\log\log \mathrm{N})})$, for $\epsilon > 0$, there is no approximation algorithm for the minimum multicolored tree partition of a $|V|$-vertex graph with its performance of $50/521(1-\epsilon)\ln|V|$. □

Corollary 4.2. *The Minimum Multicolored Tree Partition problem remains to be NP-complete for edge-colored bipartite graphs.*

Proof. It follows immediately from the observation that the graph G constructed in the proof of Theorem 4.1 is also bipartite. □

5. Concluding remarks

In this paper we have studied the complexity of finding a minimum monochromatic or multicolored subgraph such as clique, cycle, tree or path partition for an edge-colored graph G, especially when the color assignment changes to a certain threshold. We also derive an inapproximability result for the Minimum Multicolored Tree Partition problem when the edges of G are assigned by a general colors. If G is assigned a proper coloring, or G is a complete graph K_n or G is a complete multipartite graph, the complexity of these problems and similar problems in [8,16,17,20] can be further studied.

We have presented an approximation algorithm of factor $\ln m + 1$ to solve the Minimum Monochromatic or Multicolored Clique Partition problem for a K_4^--free graph with the largest monochromatic clique or multicolored clique of size m. Simple observations can result in the following remarks on the approximation factors for the Minimum Multicolored Tree Partition problem or Minimum Multicolored Path Partition problem. Denote by $|S_{OPT}|$ the optimal solution and by M and I_M a maximum matching and the set of unsaturated vertices, respectively. If we use k colors to color the edge set of a graph of order n, then the best possible case is that every tree or path in a partition contains exactly $k+1$ vertices with k edges of different colors. It follows that $n/(k+1) \leq |S_{OPT}| \leq |M| + |I_M| \leq n$. Thus, the approximation factor is at most $\frac{n}{n/(k+1)} = k+1$. If G has a perfect matching, then $n/(k+1) \leq |S_{OPT}| \leq n/2$, which means that we can get a better approximation factor $\frac{n/2}{n/(k+1)} = (k+1)/2$. A similar analysis can be applied to the Minimum Multicolored Cycle or Clique Partition problem.

Acknowledgements

The authors are indebted to Prof. Baogang Xu and Prof. Shenggui Zhang for helpful discussions and the referees for their helpful suggestions.

References

[1] M. Albert, A. Frieze, B. Reed, Multicolored hamilton cycles, Electron. J. Combin. 2 (R10) (1995) 1–13.
[2] N. Alon, Multicolored forests in bipartite decompositions of graphs, J. Combin. Theory Ser. B 53 (1991) 143–148.
[3] N. Alon, G. Ding, B. Oporowski, D. Vertigan, Partitioning into graphs with only small components, J. Combin. Theory Ser. B 87 (2003) 231–243.
[4] A. Brandstadt, Partitions of graphs into one or two independent sets and cliques, Discrete Math. 152 (1996) 47–54.
[5] H.J. Broersma, X.L. Li, G. Woeginger, S.G. Zhang, Paths and cycles in colored graphs, Australas. J. Combin. 31 (2005) 299–311.
[6] R.A. Brualdi, S. Hollingsworth, Multicolored trees in complete graphs, J. Combin. Theory Ser. B 68 (1996) 310–313.
[7] H. Enomoto, Graph partition problems into cycles and paths, Discrete Math. 233 (2001) 93–101.
[8] P. Erdős, A. Gyárfás, L. Pyber, Vertex coverings by monochromatic cycles and trees, J. Combin. Theory Ser. B 51 (1991) 90–95.
[9] P. Erdős, Z. Tuza, Rainbow subgraphs in edge-colorings of complete graphs, Ann. Discrete Math. 55 (1993) 81–88.
[10] S. Eidenbenz, C. Stamm, P. Widmayer, Positioning guards at fixed height above a terrain—an optimum inapproximability result, in: Proceedings of the 6th Annual European Symposium on Algorithms, in: Lecture Notes in Computer Science, vol. 1461, 1998, pp. 187–198.
[11] T. Feder, P. Hell, S. Klein, R. Motwani, Complexity of graph partition problems, in: Proceedings of Thirty-First Annual ACM Symposium on Theory of Computing STOC'99, 1999, pp. 464–472.
[12] U. Feige, A threshold of $\ln n$ for approximating set cover, in: Proceedings of the 28th Annual ACM Symposium on the Theory of Computing STOC'96, 1996, pp. 314–318.
[13] M.R. Garey, D.S. Johnson, Computers and Intractability, W.H. Freeman, San Francisco, 1979.
[14] J. Hastad, Clique is hard to approximate within $n^{1-\epsilon}$, Acta Math. 182 (1999) 105–142.
[15] L. Holyer, The *NP*-completeness of some edge-partition problems, SIAM J. Comput. 10 (1981) 713–717.
[16] P.E. Haxell, Y. Kohayakawa, Partitioning by monochromatic trees, J. Combin. Theory Ser. B 68 (1996) 218–222.
[17] P.E. Haxell, Partitioning complete bipartite graphs by monochromatic cycles, J. Combin. Theory Ser. B 69 (1997) 210–218.
[18] Z.M. Jin, M. Kano, X.L. Li, B. Wei, Partitioning 2-edge-colored complete multipartite graphs into monochromatic cycles, paths and trees, J. Comb. Optim. 11 (4) (2006) 445–454.
[19] Z.M. Jin, X.L. Li, The complexity for partitioning graphs by monochromatic trees, cycles and paths, Int. J. Comput. Math. 81 (11) (2004) 1357–1362.
[20] A. Kaneko, M. Kano, K. Suzuki, Partitioning complete multipartite graphs by monochromatic trees, J. Graph Theory 48 (2) (2005) 133–141.
[21] E.L. Lawler, Combinatorial Optimization: Networks and Matroids, Holt, Rinehart and Winston, New York, 1976.
[22] G. MacGillivray, M.L. Yu, Generalized partitions of graphs, Discrete Appl. Math. 91 (1999) 143–153.
[23] Y. Manoussakis, M. Spyratos, Zs. Tuza, M. Voigt, Minimal colorings for properly colored subgraphs, Graphs Combin. 12 (1996) 345–360.
[24] V. Yegnanarayanan, Graph colorings and partitions, Theoret. Comput. Sci. 263 (2001) 59–74.